Energy & Buildings 226 (2020) 110372

Contents lists available at ScienceDirect

Energy & Buildings

journal homepage: www.elsevier.com/locate/enb

Improving cooling load prediction reliability for HVAC system using Monte-Carlo simulation to deal with uncertainties in input variables

Chengliang Fan [a], Yundan Liao [a], Guang Zhou [b], Xiaoqing Zhou [a,*], Yunfei Ding [a]

[a] Academy of Building Energy Efficiency, Guangdong Provincial Key Laboratory of Building Energy Efficiency and Application Technologies, School of Civil Engineering, Guangzhou University, Guangzhou 510006, China

[b] Zhongkai University of Agriculture and Engineering, Guangzhou 510006, China

ARTICLE INFO

Article history:
Received 14 May 2020
Revised 8 July 2020
Accepted 2 August 2020
Available online 4 August 2020

Keywords:
Cooling load prediction
Uncertainty analysis
Sensitivity analysis
Monte-Carlo simulation
Support vector machine (SVM) model
HVAC system

ABSTRACT

Accurate and reliable cooling load prediction is the basis for model predictive control (MPC) of HVAC systems. Cooling load prediction process is constrained by many input variables. These input variables contain different kinds of uncertainties that should not be ignored in pursuing a reliable prediction result. However, previous load prediction methods usually improve the prediction accuracy by optimizing model structure but fail to consider the influence of input variables uncertainties. To reduce the influence of uncertainty of input variables forecast data scientifically, this paper proposes an improved cooling load prediction reliability method in which the input variables are calibrated offline with Monte-Carlo simulations and stochastic treatment before inputting them into the prediction model. A machine learning algorithm (support vector machine, SVM) is used as the prediction model. Uncertainties in weather parameters, indoor heat gains, and historical cooling load were taken into consideration. Uncertainty analysis results showed that prediction accuracy with calibrated input variables was much better (R^2 = 0.9627) than those predicted with forecast input data directly (R^2 = 0.9488), which was closer to load prediction obtained by using measured input data (R^2 = 0.9632). The coefficient of variation (CV) value using the calibrated input data was less than 15%. The reliable cooling load prediction results are beneficial for more informed, reliable decisions on MPC application, thereby promoting energy conservation potential of HVAC systems.

1. Introduction

Heating, ventilation, and air-conditioning (HVAC) systems were responsible for 40% (or even higher) of the commercial building energy consumption [1,2]. There is great potential for energy savings of HVAC system operation. However, equipment operating condition usually does not match the best performance because of a lack of supervisory control and high-quality operation management for HVAC system in many buildings, resulting in HVAC system's more energy consumption. To address these limitations, it would be preferable to supply the required cooling load and adjust the dynamic operation parameters of HVAC system in advance [3]. Model predictive control (MPC) is regarded as an optimally advanced control technology that is critical for achieving energy conservation in HVAC systems [4,5]. Thus, accurate and reliable cooling load prediction results are the basis for MPC. Among the input variables of the load prediction model, weather changes and internal heat disturbances play a very important role in cooling load prediction, which eventually affects the MPC performance and building energy consumption.

Many methods developed for predicting the cooling load can be broadly divided into regression analysis [6,7], time series [8,9], artificial intelligence (AI) [10–14], and simulation software [15,16]. Meanwhile, many studies have focused on improving prediction accuracy via optimizing the model structure or developing the calibration method. Sun et al. [17] enhanced R^2 from 0.89 to 0.96 with a calibration method for a simplified load prediction model. Guo et al. [18] developed the MLR model with a dynamic two-step correction, and mean absolute relative error was less than 8%. Guo et al. [19] optimized the ARX model with two-stage weighted least squares regression. Li et al. [20] introduced an online forecasting model that can achieve R^2 over 0.95 forecasting accuracy. Most of the load prediction methods are a fixed-point process, which means input variables forecast data are fixed value, as is the cooling load.

However, the traditional load prediction methods usually improve the prediction accuracy by modulating the model struc-

* Corresponding author.
E-mail address: zhouxq@gzhu.edu.cn (X. Zhou).

https://doi.org/10.1016/j.enbuild.2020.110372
0378-7788/

ture or prediction logic, but less consider the uncertainty influence from input variables. Uncertainty refers to a state of having limited knowledge where it is impossible to exactly describe the state [21] and exists widely in engineering processes, such as in the prediction process, modeling, control, and management [22]. Cooling load demand is constrained randomly by many factors at the next moment, such as HVAC system operating characteristics (e.g., operation schemes variation and control deviation), weather changes (e.g., outdoor dry-bulb temperature and solar radiation) and internal heat disturbances (e.g., occupancy and lighting power). During the forecasting process, these factors were combined into the basic data of the load prediction model, and it was assumed that the weather forecast was 100% accurate and without any uncertainty [23]. These input variables forecast data are usually difficult to guarantee complete accuracy and reliability in the practice prediction process. Hence, the predicted input variables inevitably exit deviation. These uncertainties then follow the input variables into the prediction model and ultimately lead to deviations between the predicted load and the actual measured load. Potential uncertainty in the prediction process may cause the predicted load to be unrealistic, which could lead to the misfunctioning of the system and a deterioration in the performance of the chiller plant [24,25]. Therefore, there is inherent uncertainty in the forecast input variables, which cannot be neglected when conducting prediction.

Recognizing the importance of the influence of uncertainty in cooling load prediction and operation of HVAC systems, uncertainty analysis has gradually been used in HVAC system simulation [26–28] and heat transfer calculation [29]. Uncertainty sources are measurement error, prediction model uncertainty, and predicted input variables uncertainty [30]. For example, Liao [31,32] and Liu et al. [33] studied uncertainty of chillers sequencing control. Parisio et al. [34] illustrated how HVAC control strategies considered several sources of uncertainties, e.g., occupancy and weather conditions. Massoumy et al. [35] introduced a method to handle model uncertainty in HVAC control processes. Huang et al. [36] investigated uncertainties that existed in the chiller plant capacity and developed a method to identify them. Zhao et al. [37] analyzed the uncertainty of weather forecasted data had a great influence on prediction accuracy. The mean absolute percentage error of load prediction was reduced from 11.54% to 10.92% based on uncertainty analysis.

Overviewing the previous researches, although recent literature demonstrates that uncertainty may have an important impact on HVAC operation control and cooling load prediction, there are still few studies to analyze the influence of uncertainty of input variables (e.g., weather parameters, indoor heat gains, and historical cooling load) on the load prediction results. However, weather parameters and indoor heat gains play a significant role in the dynamic cooling loads of a building. If the uncertainty of these input variables is ignored, it may cause errors in prediction model inputs, which eventually reduces the accuracy and reliability of model output cooling load prediction results. These uncertainties of input variables also would cause the predicted cooling load may vary within a range rather than a certain or fixed value. In order to improve the accuracy and reliability of cooling load prediction results, this paper quantified and reduced the impact of uncertainties of input variables forecast data on load prediction results using Monte-Carlo (MC) method. Firstly, the forecasting error was obtained by comparing input variables forecast data with measured data. The error characteristic was described with a two-periods probability distribution. Secondly, the error sampling results of input variables forecast data were estimated based on MC method. Thirdly, the input variables forecast data was calibrated offline with error sampling results. Hence, the uncertainties of input variables forecast data were reduced after the calibrating process. Finally, the hourly cooling load with the highest probability was selected as the prediction load. Support vector machine (SVM), a machine learning method based on statistical theory, is selected as the prediction model in this paper.

The rest of paper is organized as follows: Section 2 presents the principles of MC simulation method and calibration method of input variables forecast data based on MC simulation. Section 3 presents the sensitivity analysis method and SVM model. Section 4 identifies the main input variables contributing to the load prediction through a sensitivity analysis, and a case study is used to illustrate how the methodology can be applied to quantify and reduce the impact of uncertainty of the input variables forecast on load prediction results. Section 5 shows the results and discussion, as well as some proposals for future work. Section 6 presents conclusions where the main results are emphasized.

2. Methodology

The basic idea of analyzing input variables uncertainty to improve the reliability of cooling load prediction is illustrated in Fig. 1.

1) The dominant variables affecting the dynamic cooling load are identified based on sensitivity analysis and then used as input variables to construct the prediction model.
2) The forecast input variables are treated as dynamic variables. Errors between forecast input variables and measured data are calculated. Those errors are described with specific probability distributions. With this method, the uncertainties of forecast input variables are quantified.
3) Error samplings are randomly generated using probability distribution models. Each forecast input variable is calibrated with error sampling using MC simulation. Then, the calibrated forecast input variables are input to the prediction model.
4) Finally, the predicted cooling load is calculated by the prediction model. The load with the largest probability is taken as the prediction load at the next moment.

2.1. Uncertainty analysis

The cooling load prediction process is constrained by many variables (e.g., HVAC system operating characteristics, weather changes, building envelope performance variations, and internal heat disturbances randomly) as described in the previous studies [3,14]. These variables are clarified into two types, i.e., design variables and measured variables. Design variables, such as occupancy (*Occ*), lighting power (*LP*), equipment power (*EP*), infiltration rate (*IR*) and ventilation rate (*VR*), have uncertainties because some of them are related to human behavior (e.g., *Occ*, *LP* and *EP* are mainly affected by human custom subjectively), or some of them cannot be set to match the design in practical operation (e.g., *IR* and *VR* cannot be measured accurately). Measured variables refer to the variables that are relative to the real-time measurements of affecting building cooling load during HVAC operation, including weather variables (e.g., outdoor dry-bulb temperature (*T*), relative humidity (*RH*), wind speed (*WS*), solar radiation (*SR*)) and historical cooling load (CL_a). Weather variables have uncertainties because they are neither predictable accurately nor controlled at the design stage. The process of predicting input variables inevitably exits deviation.

Therefore, the inherent uncertainty is widespread in the input variables forecast data and cannot be ignored when performing prediction. Those uncertainties would be taken into the prediction model along with input variables and finally cause deviations in predicted loads. Uncertainty analysis aims to make a technical con-

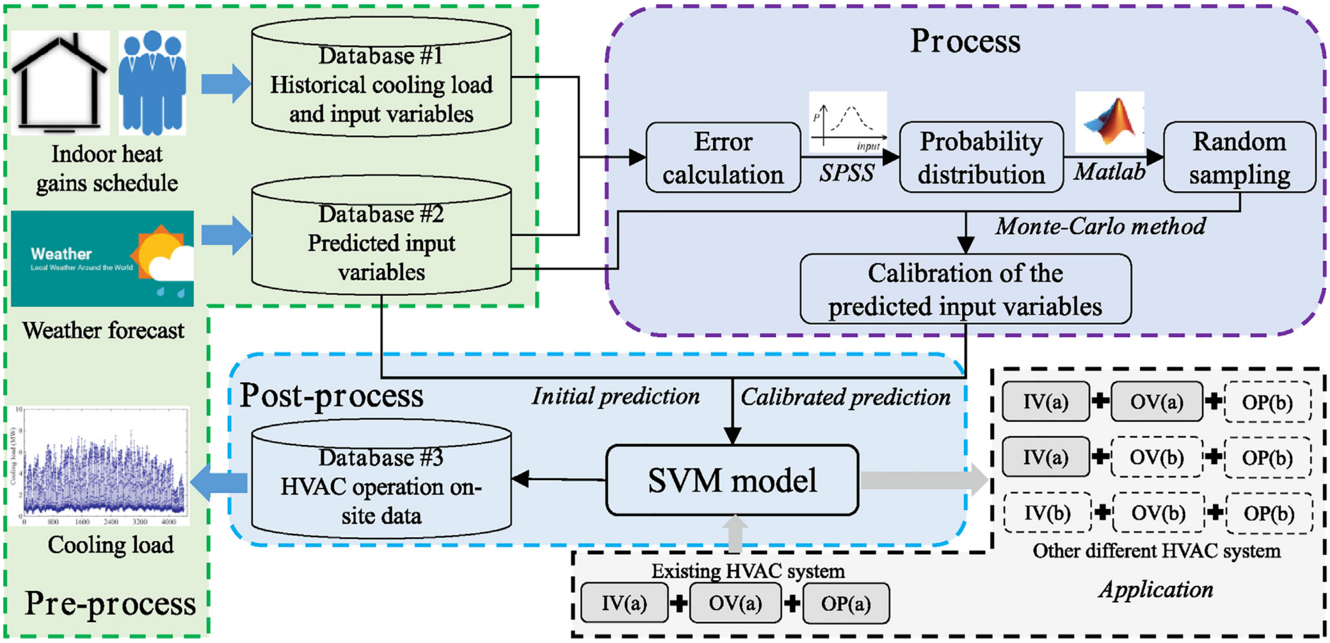

Fig. 1. The workflow of cooling load prediction.

tribution to quantify uncertainties of relevant input variables. In this study, uncertainties of weather parameters, indoor heat gains, and historical cooling load are taken into consideration. In general, uncertain variables can be quantified to follow the normal distribution, triangular distribution, and uniform distribution. The probability distribution of input variables can be modeled directly referencing to available databases or historical on-site measurement information. For example, the uncertainties in weather variables are described using a normal distribution from existent data [38].

2.2. Monte-Carlo simulations

When the variable or parameter itself has probabilistic characteristics, the sampling results can be generated by computer simulation methods based on Monte-Carlo (MC) method. The statistics or parameter values can be calculated based on those samplings [37]. In this study, MC simulation is used to calibrate the forecast input variables with the prior error probability distributions. The implementation of the MC simulation mainly contains three steps [38]: describe the random error probability distribution of variables forecast data, obtain sampling results from the specific probability distribution, and analyze cooling load prediction statistical results. Those steps are illustrated in Fig. 2. Firstly, the input variable uncertainty can be quantified by calculating the error probability distribution of input variables forecast data. Then, the sampling results for each input variable are generated with the specific probability distributions using a random sampling method. The characteristic of distribution can be revealed using this random sampling process at each time, which repeats n times if n samples are required. Sampling size n is considered as variables uncertainty range to study their effect on the cooling load prediction accuracy.

To perform MC simulation for i input variables, a row of samples $\{x_{11}\ x_{12}\ \ldots x_{1i}\}$ was generated from the specified probability distribution, and all these rows form an input matrix. The input matrix X was then imported into the prediction model $f(\cdot)$ for forecasting the cooling load Y. The number of iterations was calculated n times, producing a set of possible results that were used to represent the effect of input variables uncertainties on the output prediction cooling load. One hour ahead time interval can characterize cooling load patterns parsimoniously, which is adequate for the prediction process [39]. Hence, the time interval of the input variables is one hour in this paper.

2.3. Calibrate the predicted input variables based on MC method

The initial cooling load prediction result Y_{initial} can be calculated directly using the predicted input variables and the historical data, as given by Eq. (1).

$$Y_{initial} = f(X_t, CL_{t-1}) \tag{1}$$

X_t is the predicted input matrix at time t, $f(\cdot)$ is the prediction model. $CL_{t\text{-}1}$ means the previously measured load data, which proposes the following special rules: 1) During daytime (i.e., 9:00–18:00), $CL_{t\text{-}1}$ should be the 1-h ahead measured data; 2) For the night time (i.e., 19:00–8:00), $CL_{t\text{-}1}$ presents measured data over the same time on the reference day. The occupancy schedule and non-occupancy period both have great impacts on the initial thermal status of buildings. Days with similar occupancy and non-occupancy schedules are classified as the same classification, and the nearest day is selected as the reference day for simplification [3,17]. The classification results for a given reference day are shown in Table 1.

As described in Fig. 1, MC simulation is used to calibrate the predicted input variables with measured error probability distributions. Firstly, the prior prediction error (θ) is calculated first using the difference between historical input variables forecast data (X) and measured data (X^*). Then, the specific error probability distribution characteristics can be described as $P(\theta|X^*)$. Meanwhile, the program of Monte-Carlo simulation is developed in MATLAB to implement the random sampling of its error. Finally, each forecast input variable (X) is calibrated with the predicted error sampling results ($\Delta\theta$). The calibrated input variables $X_t|_{P(\theta|X^*)}$ can be defined as follows:

$$X_t|_{P(\theta|X^*)} = X_t + \Delta\theta \tag{2}$$

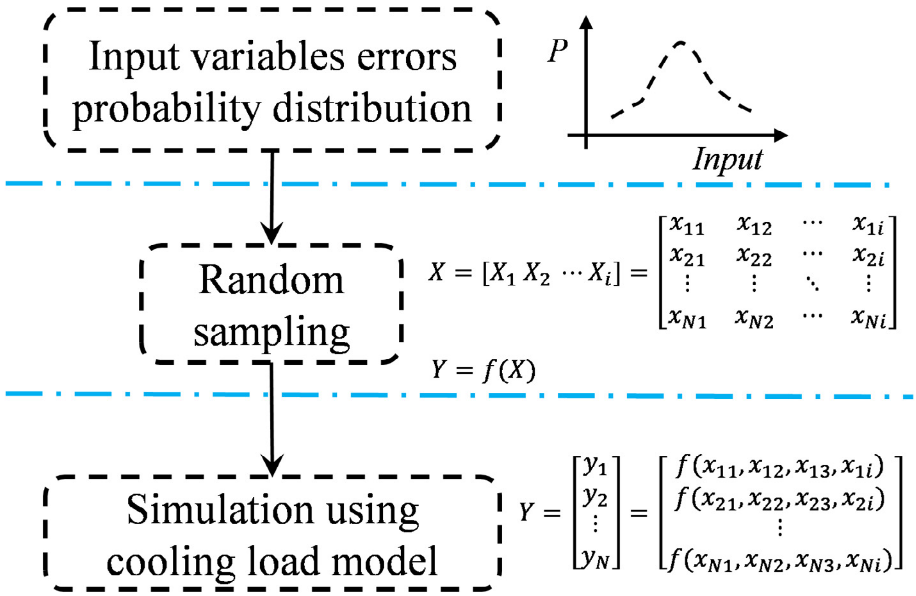

Fig. 2. Process of MC simulation.

Table 1
Reference day selection classification.

Day types	Day	Reference day
Weekdays	Monday/day after the holiday	Last Monday
	Tuesday	Last Tuesday
	Wednesday-Friday	Previous day
Holidays	Weekends/holiday	Last Saturday
Events days	Special days when events occur	Last events day

$X_t|_{P(\theta|X^*)}$ means the forecast input matrix X_t calibrated by the prior error probability $P(\theta|X^*)$. $\Delta\theta$ is the predicted error, which is generated from error probability distribution $P(\theta|X^*)$ based on MC simulation results. After finishing MC simulation, the initial prediction load Y_{initial} is calibrated as Y_{calib}, as given by Eq. (3).

$$Y_{calib} = f\left(X_t|_{P(\theta|X^*)}, CL_{t-1}\right) \tag{3}$$

To reduce computational resources and save the prediction time, the calibration process (i.e., Eq. (2)) and prediction process (i.e., Eq. (3)) are divided into two processes, offline simulation and online application, respectively. The offline simulation includes prior error probability description and error sampling based on MC method. After finishing the MC simulation, the online application directly uses the prediction errors that have been generated offline to calibrate the forecast variables. To facilitate the application, the available input data (i.e., forecast input variables and historical cooling load) is gathered to finish the online load prediction during HVAC operation.

As the cooling load prediction process is similar in different HVAC systems with some common input variables or even similar in all HVAC systems, the calibrated methods based on the variables uncertainties can be used as inputs to other HVAC or as a reference to optimize a new HVAC system operation. Fig. 1 illustrates the application limitation condition of the proposed calibration method for the different HVAC systems. The operational cooling load prediction of a given HVAC system is determined by these types of variables - namely indoor variables (IV), outdoor variables (OV), and operation parameters (OP). By categorizing HVAC systems with common characteristics into different groups (e.g., groups (a) and (b)), the generic variables for a specific group can be determined. For example, the HVAC system (a) is determined by the combination of $IV_{(a)}$, $OV_{(a)}$, and $OP_{(a)}$, the HVAC system (b) is determined by the combination of $IV_{(a)}$, $OV_{(a)}$, and $OP_{(b)}$. $OP_{(a)}$ and $OP_{(b)}$ are different parameter groups. The identification of the set of generic variables in different categories of HVAC systems will make the prediction model and calibration method apply to more scenarios.

3. The load prediction model using SVM

Building cooling load has typically nonlinear and dynamic characteristics. One of the data-driven models (i.e., SVM model) is selected as the prediction model. SVM has a strong generalization ability to effectively solve practical problems such as small samples and nonlinearities. It is suitable to handle nonlinear and dynamic characteristics, such as cooling load. Besides, its efficiency and accuracy are better than many models.

3.1. Pre-process of input variables data

All input variables have different sizes and units, the prediction process will not launch if the input data is not pre-processed. So, it is necessary to normalize all variables within the same range, e.g., [−1,1] or [0,1], and all input variables data can then be refined to dimensionless values. After the prediction process is completed, an anti-normalized process may be conducted to restore the original cooling load value. The normalized function x_i' and the anti-normalized function x_i can be expressed as Eqs. (4) and (5).

$$x_i' = 2 \times \frac{x_i - x_{min}}{x_{max} - x_{min}} - 1 \tag{4}$$

$$x_i = \frac{1}{2} \times (x_i' + 1) \times (x_{max} - x_{min}) + x_{min} \tag{5}$$

here, x_{max} and x_{min} represent the maximum and minimum values of input variable x, respectively.

3.2. Selecting input variables with sensitivity analysis

The cooling load prediction results can be usually affected by many factors (e.g., HVAC system operating characteristics, weather changes, building envelope parameters, and internal heat gains). Input variables of the model with high sensitivity for the cooling load would be considered. For this research, cooling load prediction is conducted on a specific building and HVAC system. Building envelope parameters and HVAC operation factors are stable during the analysis period. Thus, weather changes and indoor heat gains are dominating factors for predicting the dynamic cooling load of HVAC system operation. Sensitivity analysis is used to study the mapping relations of input variables and output cooling load. There are lots of methods among previous studies to conduct sensitivity analysis, e.g., Pearson correlation coefficient [40], standardized regression coefficient [41], and sensitivity coefficient [42]. A simple sensitivity analysis method, which directly reflects the cooling load benefit influence by changing relevant input variables, has been adopted as the sensitivity quantification indicator (SQ) in this study. As shown in Eq. (6), SQ can explain the load affected by these factors.

$$SQ_j = ACL_{interval,j} - ACL_{base,j} \tag{6}$$

where j presents the kind of input variables, $ACL_{interval}$ is the cooling load by changing the relevant input variables with a specific interval, e.g., the 1*interval for outdoor dry-bulb temperature is 1 °C each time, ACL_{base} is the baseline cooling load by using the design data of the building.

3.3. Support vector machine (SVM) model

As described in Section 3.2, the high sensitivity variables are selected as inputs in the prediction model. Then, the support vector regression (SVR) model, which belongs to SVM [43], is used to predict the cooling load. Suppose that all the input variables X_i and outputs Y_i samples are defined as$\{(X_i, Y_i)\}_{i=1}^{n}$. In the method of SVM for regression, the nonlinear problem is transformed into a linear problem, which also is a low-dimensional space transformed into a high-dimensional feature space by a nonlinear mapping. After transforming, the relationship between the outputs and inputs can be described as Eq. (7). $\varphi(X)$ represents nonlinearly mapped from the input space X , and coefficients W and b were estimated by minimizing the regularized risk function using Eq. (8).

$$f(X) = W \cdot \varphi(X) + b \tag{7}$$

$$\begin{cases} Minimize: \frac{1}{2}\|W\|^2 + C\frac{1}{n}\sum_{i=1}^{n} L(Y_i, f(X_i)) \\ L(Y_i, f(X_i)) = \begin{cases} 0, |Y_i - f(X_i)| \le \varepsilon \\ |Y_i - f(X_i)| - \varepsilon, \quad \text{others} \end{cases} \end{cases} \tag{8}$$

Here, $||W||^2$ represents the regularized term, $L(\cdot)$ is the empirical error measured by the insensitive loss function. Y_i is the actual output result of node i. The insensitive loss function ε represents the width of the confidence interval. C is the penalty parameter. To get the estimations of W and b, positive slack variables ζ_i and ζ_i^* were introduced, then Eq. (8) was transformed to Eq. (9).

$$\begin{cases} Minimize: \frac{1}{2}\|W\|^2 + C\frac{1}{n}\sum_{i=1}^{n}(\zeta_i + \zeta_i^*) \\ \text{subject to}: Y_i - f(X_i) \le \varepsilon + \zeta_i \\ f(X_i) - Y_i \le \varepsilon + \zeta_i^* \\ i = 1, 2, \cdots N; \zeta_i, \zeta_i^* \ge 0 \end{cases} \tag{9}$$

Lagrangian multipliers α_i and α_i^* were introduced to obtain the dual form of the original objective function [44], as shown in Eq. (10). Only those corresponding to (α_i-α_i^*) ≠ 0 were included (i.e., support vectors $X_i \in SV$). Moreover, b can also be computed using Eq. (12).

$$f(X) = \sum_{i=1}^{n}(\alpha_i - \alpha_i^*)K(X_i, X) + b \tag{10}$$

$$K(X_i, X) = \varphi(X_i) \cdot \varphi(X) \tag{11}$$

$$b = \frac{1}{N_1}\left\{ \begin{array}{l} \sum_{\alpha_i \epsilon (0,c)} \left[Y_i - \sum_{X_j \epsilon SV} \left(\alpha_j - \alpha_j^* \right) K(X_i, X_j) - \varepsilon \right] \\ + \sum_{\alpha_i \epsilon (0,c)} \left[Y_i - \sum_{X_j \epsilon SV} \left(\alpha_j - \alpha_j^* \right) K(X_i, X_j) + \varepsilon \right] \end{array} \right\} \tag{12}$$

N_1 is the number of support vectors with either {$\alpha_i \in$ (0,C), α_i^* =0} or {$\alpha_i^* \in$ (0,C), α_i = 0}. $K(X_i, X)$ is the kernel function, which affects the performance of the SVR model. In this study, the Gaussian function can be used for $K(X_i, X)$ as showing Eq. (13). γ is the parameter of the kernel function. For SVR model, the prediction accuracy depends on the optimal combination of three parameters (C, ε, and γ), which can be optimized by the grid search algorithm. As MC simulation performed offline, the cooling load model run times depend on the speed of the online grid search algorithm.

$$K(X_i, X) = \exp(-\gamma\|X_i - X\|^2) \tag{13}$$

3.4. Evaluation indexes

The characteristics of the model output load can be investigated using statistical evaluation indices. The coefficient of determination (R^2), root mean squared error (RMSE), coefficient of variation (CV) and mean absolute percentage error (MAPE) are used to evaluate the accuracy of prediction results. These evaluation indexes can be defined and calculated as follows. R^2 is used to evaluate how well the predicted load data matches the actual data. The higher R^2 value, the greater the model prediction performance. For most engineering purposes, a regression model can be accepted when R^2 > 0.7 [40]. RMSE is used to amplify the variance in data. MAPE is a measure of prediction accuracy, which considers not only the error between the predicted and the measured values but also the ratio between them. Previous studies have specified that CV below 30% is sufficient for engineering applications [45].

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(CL_{a,i} - CL_{p,i})^2}{\sum_{i=1}^{n}(CL_{a,i} - CL_{a,mean})^2} \tag{14}$$

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}(CL_{a,i} - CL_{p,i})^2}{n}} \tag{15}$$

$$\text{CV} = \frac{RMSE}{CL_{a,mean}} \times 100\% \tag{16}$$

$$\text{MAPE} = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{CL_{a,i} - CL_{p,i}}{CL_{a,i}}\right| \times 100\% \tag{17}$$

where CL_a is the actual cooling load, CL_p is the predicted cooling load, $CL_{a,mean}$ is the mean value of the actual load.

4. Case studies

4.1. Building description and on-site measurement

For case studies, a library building with a construction area of 98178 m^2 located in Guangzhou was selected. A real HVAC system is served by a district cooling system that provides chilled water to the HVAC system. The library was closed on Wednesday (close day) and open on other days (open day). The HVAC system operation scheme was based on these two days types. The operation set of this HVAC system is listed in Table 2. For the 9th floor of the North building, it is open 24/7. For the rest of the floors, they are open from 7:30 to 20:30.

Table 2
HVAC system operation schedule in different zones.

Zones	System operation period	
	Open day	Close day
South building	7:30–17:00	–
North building (except 9th)	7:30–20:30	–
9th floor of North building	0:00–24:00	0:00–24:00

The input variables of the prediction model and cooling load data of the HVAC system were hourly measured with test instruments and the building's automatic control system. Changes in occupancy rate and light power were related to the library's daily actual schedule. The on-site measurement period was from June 21 to July 11, July 16 to July 22, 2018. Fig. 3 shows measured input variables for one week. The on-site measured cooling load data is shown in Fig. 4.

4.2. Sensitivity analysis results

The intervals for each variable reference to the previous study [38] when conducting sensitivity analysis. Table 3 summarizes the different interval types. Weather factors include outdoor dry-bulb temperature (*T*), outdoor relative humidity (*RH*), and solar radiation (*SR*). For refining the effect of indoor heat gains on cooling load, lighting power density (*LP*) and occupancy density (*Occ*), were taken into consideration. The sensitivity was analyzed based on China standard weather data (CSWD). EnergyPlus was used to perform the simulation. Variables were adjusted to fit EnergyPlus parameter format. For sensitivity results of input variables to cooling load predictions, the cooling load difference due to 1*interval changes in input parameters is the one-year total cooling load difference. The sensitivity analysis results are shown in Fig. 5, the greater difference in one-year total cooling load, the greater sensitivity of the changing factor for output cooling load. When variables changing within 1*Interval, the one-year total load difference of T was $SQ = 1.03 \times 10^6$kW, followed by RH ($SQ = 0.46 \times 10^6$kW). It means T had the larger sensitivity for the output cooling load than RH. According to SQ values, the variables with the highest sensitivity for cooling load are selected as input variables. Therefore, five variables (*T*, *RH*, *LP*, *SR*, and *Occ*) and previous 1-h measured cooling load $CL_{a,t}$ were selected as inputs for SVM model.

4.3. Modeling process of SVM model

4.3.1. Database establishment

SVM model belongs to one of the data-driven models, which are built based on sampling data. Thus, the database used for modeling should be established first. The database includes weather data, occupancy schedule, lighting power density, and historical cooling load data. In this study, actual weather data and measured load data were obtained as described in Section 4.1. The cooling load data for training was obtained by simulation due to the limited available on-site measurements. The reason why choosing the simulation data as a part of the database is as follows: firstly, the hourly on-site cooling load data has some limitations because of the difficulty in on-site continuous test during the whole air-conditioning period, a limited amount of measured data can be used to validate the proposed method; secondly, the hourly on-site cooling load data exists the measurement instrument error, but EnergyPlus simulations can make the data more integrated and accurate; thirdly, this paper focuses on method research for uncertainty analysis, the training process of the prediction model using simulation database has less influence on the uncertainty analysis results. If the on-site measurements can be easily obtained or more accurate than simulations for other building cases, measurement data also can be selected as a database for both training and testing [42]. Thus, the database can be established based on two parts of simulation results and on-site measurements, as

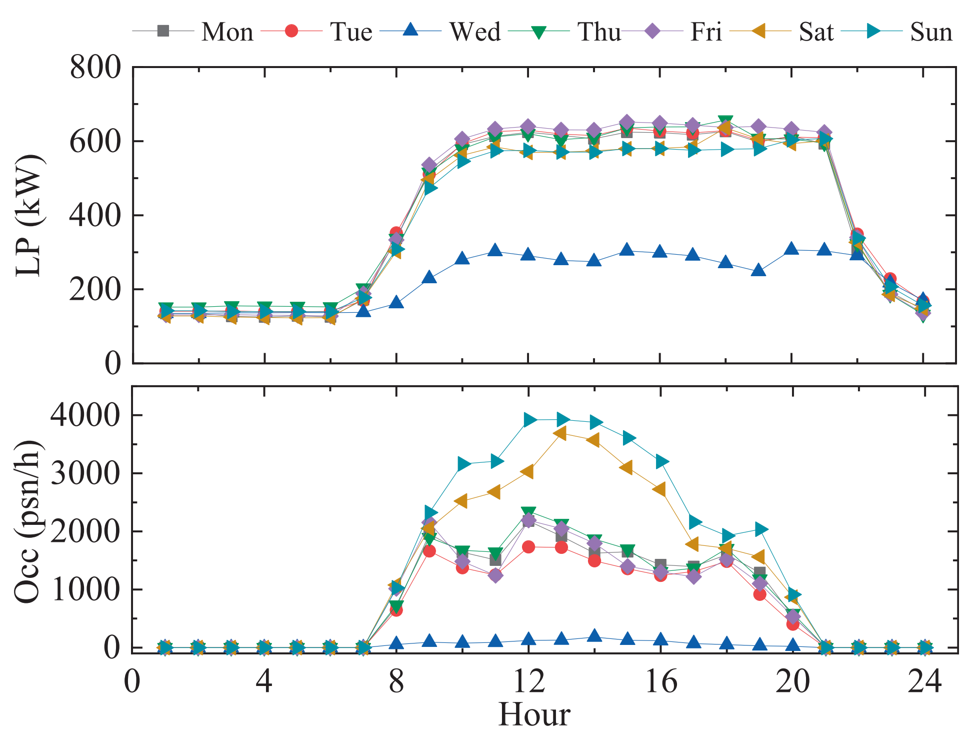

Fig. 3 (*continued*)

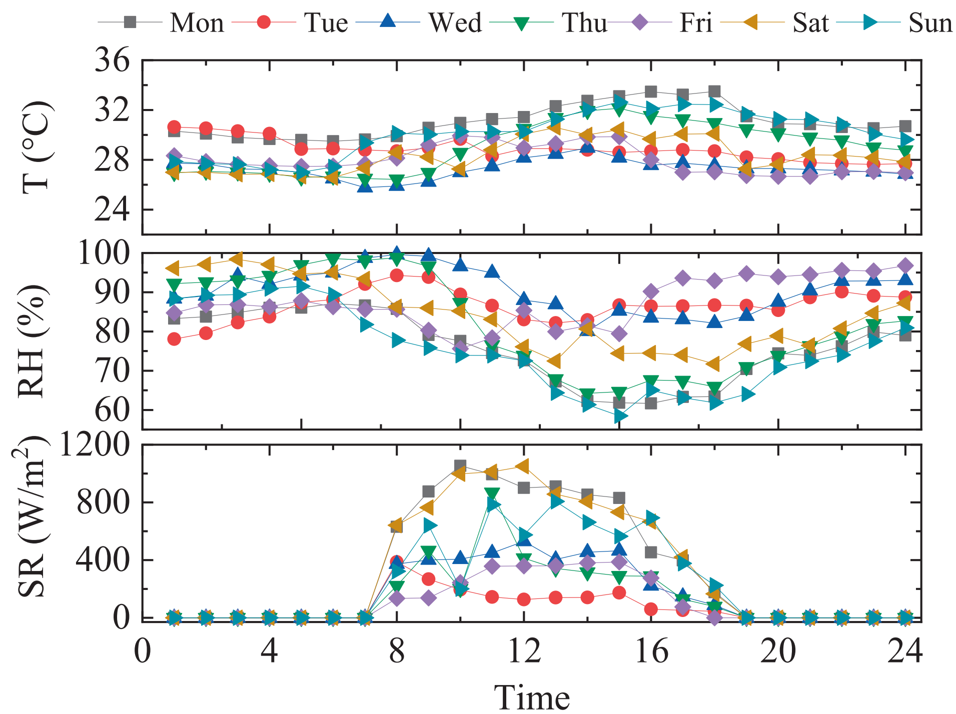

Fig. 3. Measurements of the input variables.

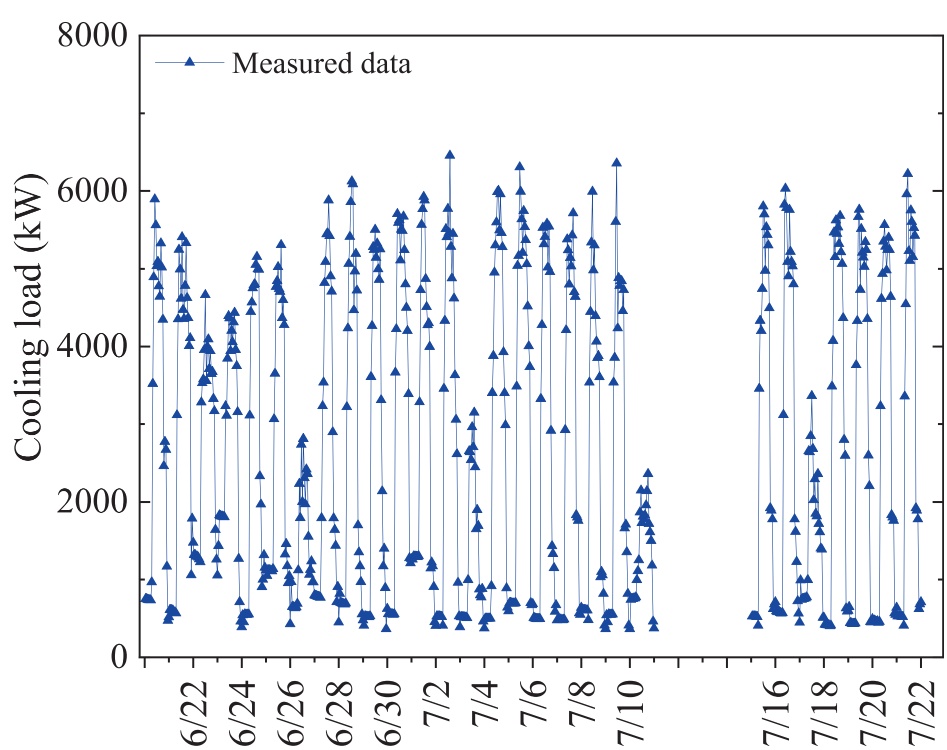

Fig. 4. On-site measured cooling load.

shown in Fig. 6. This paper used the hourly simulated cooling load from May 1 to October 31 as training samples (4416 observation data) and used actual meteorological data and the hourly measured load from on-site measurement period as test sample (672 observation data).

4.3.2. SVM model training and validations

As described in Section 3.3, one of the SVM models was built as the main model for cooling load prediction. For the proposed model, three key parameters to be optimized were the penalty parameter C, the loss function parameter ε, and the kernel function parameter γ.

Table 3
Interval set for different variables.

Variables	Unit	Baseline	Interval
Outdoor dry-bulb temperature	°C	28	1
Relative humidity	%	70	5
Solar radiation	W/m^2	350	50
Occupancy density	m^2/ psn	10	1
Lighting density	W/m^2	20	1

The parameter C, as a regularization constant, affected the structural complexity of the model. A larger C would increase the possibility of the model overfitting and training error, and vice versa. The parameter ε influenced the model decision boundary. The larger ε provided a more flexible generalization ability. The parameter γ could indicate the correlation between the support vector. Three kernel functions (i.e., Gaussian, linear, and polynomial) from the most suitable decision was chosen at different application situation. Gaussian function with the advantage of computationally efficiency, reliability, and ease of adaption for optimization, was adopted to handle the complex nonlinear relationship between the inputs and output [42]. Three key parameters of SVM model were set as $\varepsilon = 0.01, C, \gamma \in [2^{-5}, 2^5]$. The best parameter combination C and γ of the SVM model were dynamically optimized by a grid search algorithm for each time step. The computer used for the simulation was configured as a 64-bit operating system, 16 GB RAM, Intel(R) Cor(TM) i7-4790CPU 3.6 GHz. The above parameter optimization method and settings were used throughout the research. The prediction load was then obtained by using the above optimization results.

The optimization process of parameter combination C and γ is shown in Fig. 7. The best C and γ were 8 and 0.125 respectively. Fig. 8 shows the prediction performance, which presents a good fitting effect compared with the on-site load. The training and testing evaluation results are summarized in Table 4. P1 is the prediction load adopting the measured input variables. Table 4 narrates the initial prediction accuracy R^2 value is nearly 0.963. CV value is 14.6%, which means that the SVM model is sufficiently accurate for engineering applications.

5. Results and discussion

5.1. Probability distribution of input variable errors

As uncertainty described in Section 2.1, the uncertainties of each input variable belonged to a specific probability distribution referencing to previous studies. For example, a normal distribution was used to describe the uncertainties in T and RH [38]. The probability distribution of the error should be determined according to the actual error changing characteristics rather than the fixed distribution for simplification. A total of 672 samples was used to analyze the error characteristic between the measured data with the predicted data. It is found that the errors obey the normal distribution. The analysis results also show that the changes in on-site measured weather data and indoor variables had a strong regularity. For example, the outdoor temperature usually raised from 6:00 AM to 4:00 PM and decreased from 4:00 PM to 6:00 AM. If only a fixed probability distribution was used to describe this changing feature, the description may be incomplete.

Therefore, to reflect the actual temperature fluctuation, two probability distributions are used to model the error uncertainties distribution, as are the other input variables. For variables T and RH, two-periods of probability distributions N_1 and N_2 are 6:00 AM-4:00 PM and 4:00 PM-6:00 AM, respectively. For SR, two-periods of N_1 and N_2 are 8:00 AM-3:00 PM and 3:00 PM-6:00 PM. For Occ, two-periods of N_1 and N_2 are 8:00 AM-1:00 PM and 1:00 PM-9:00 PM. For LP, two-periods of N_1 and N_2 are 6:00 AM-1:00 PM and 1:00 PM-6:00 AM. SPSS statistic software and MATLAB software provide a convenient way to perform parameters probability distribution and MC simulation. Table 5 lists the uncertainty probability distributions of these input variables. The truncation was applied to all these input variables considering the practical application, such as Occ, SR, and LP. Fig. 9 shows the error probability distribution N_1 between the predicted temperature T and the measured temperature T^* at the previous time $t-1$. The mean value is 0.33 °C and the standard deviation was 0.838 °C.

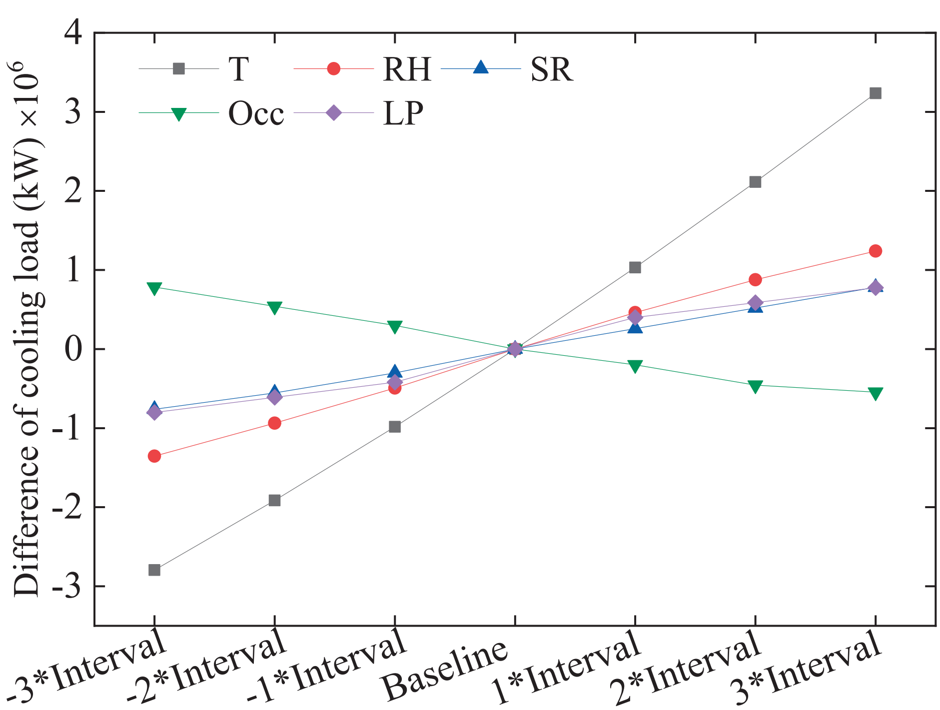

Fig. 5. Sensitivity analysis results.

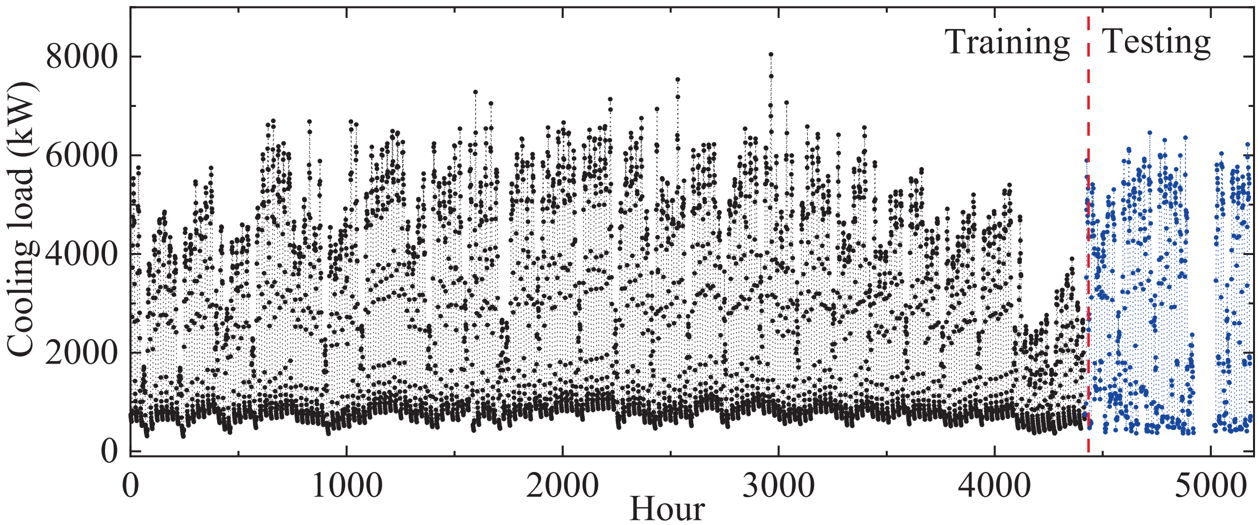

Fig. 6. Database for training and testing process.

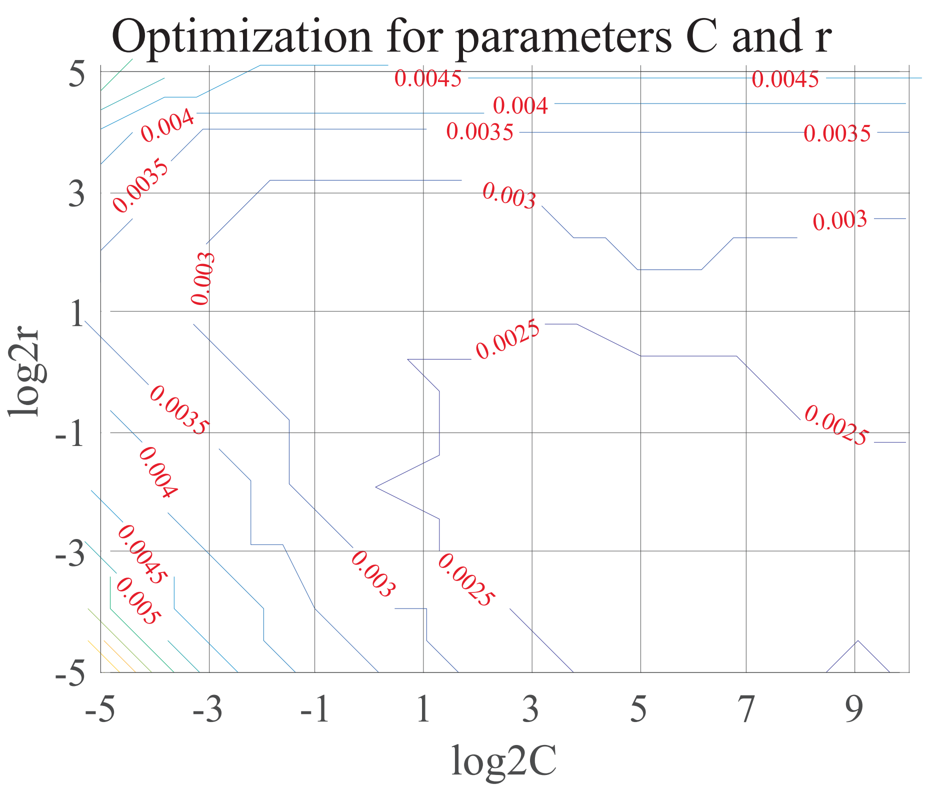

Fig. 7. Optimization for parameters C and γ.

5.2. Calibration of forecast input variables based on MC simulation

To reduce the effect of uncertainty of input variables on cooling load prediction accuracy, the MC method was used to modify the input variables data of the SVM model. Then, the calibrated input variables (i.e., weather data, and indoor heat gains data) were imported into the model for load prediction. 1-h (11:00 AM, July 17, 2018) was randomly chosen to investigate the effect of input variables uncertainty on prediction results. For each uncertain variable, a row of 10,000 samples was generated from the specified probability distribution using random sampling. Fig. 10 shows the calibration process of different input variables at time t based on MC simulation. The load prediction distribution result was obtained by using the calibrated input variables. As shown in Fig. 10, the left vertical axis is the frequency counts that a level of the variable value occurs, while the right one is the probability density of the variable value, and the horizontal axis displays the range of the input variable value. The values of mean, variance, and truncation boundaries set were calculated and calibrated using design values and real measurements.

The comparison of predicted input data and calibrated input data is shown in Table 6. P2 is the prediction load adopting input variables forecast data. P3 is the prediction load adopting the calibrated input variables based on the MC method. The predicted

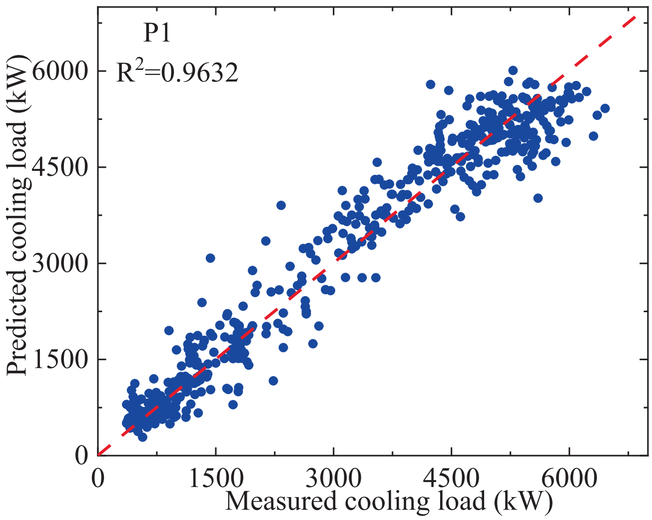

Fig. 8. The prediction results using on-site measured input data.

Table 4
Training and testing performance evaluation.

Input data type	R^2	RMSE (kW)	CV (%)	MAPE (%)
Training	0.9766	290.3	10.5	7.7
P1	0.9632	378.8	14.6	14.3

variable T(t) was 31.0 °C. The most frequent error value $\Delta\theta$ of T(t) in the results of random sampling simulation was near 0.5 °C, and the revised forecast data $T(t)^*$ was 31.5 °C through calculation, which means that the expected value of T(t) was 31.5 °C and closer to the real weather data, i.e., 32 °C. With this MC method, all 5 sample vectors (T, RH, SR, Occ, LP) and CL_a formed an input matrix

Table 5
Input variables and their error probability description.

Variables	Unit	Error probability distribution	
		N_1	N_2
T	°C	N (0.33, 0.84^2)	N (−0.23, 0.73^2)
RH	%	N (−1.77, 4.27^2)	N (1.17, 3.2^2)
SR	W/m^2	N (39.95, 203.86^2)	N (−104.53, 131.38^2)
Occ	psn/h	N (112.4, 445.83^2)	N (−146.85, 402.89^2)
LP	kW	N (30.18, 40.4^2)	N (−7.13, 16.94^2)

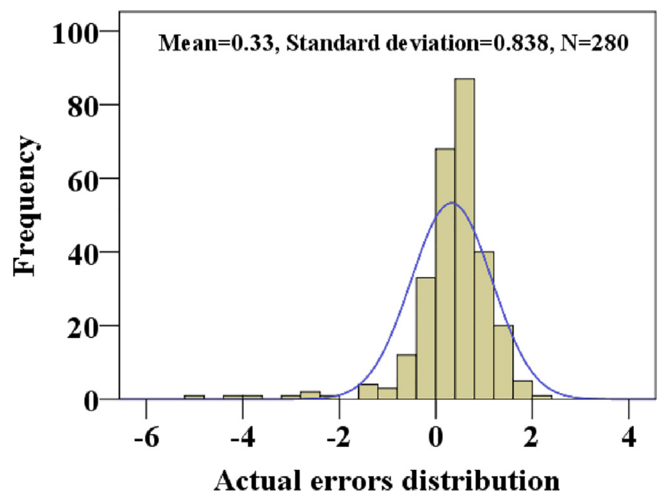

Fig. 9. Example of error of probability distribution of variable T(t − 1).

X. Then, X was imported into the cooling load prediction model. To visually understand uncertainty, uncertainty should be converted to a cooling load index. As shown in Table 6, it can be found that the predicted load P3 was 5692.5 kW, closer to measured load

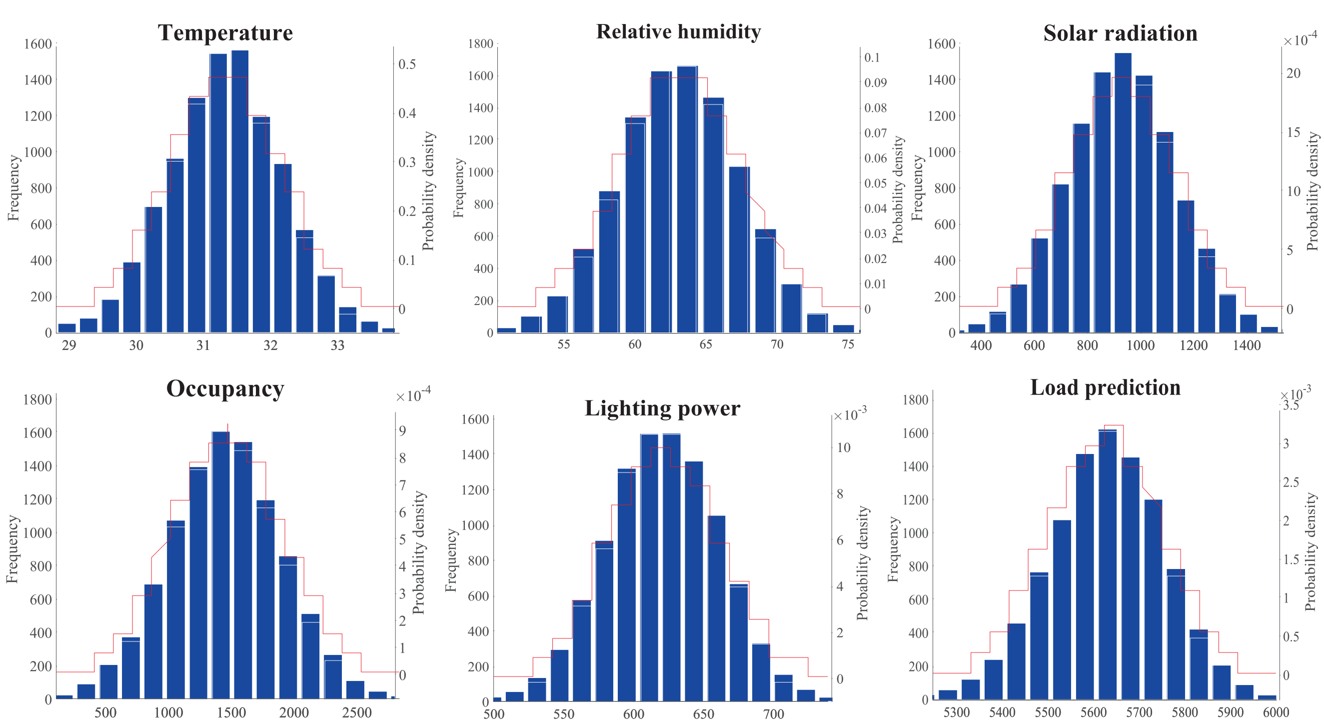

Fig. 10. The MC calibration process at time t.

Table 6
Load prediction results using the different input variables type.

Input data type	*T*(t) °C	*RH*(t) %	*SR*(t) W/m^2	*Occ*(t) psn/h	*LP*(t) kW	CL_a(t − 1) kW	CL_p(t) kW
P1	32.0	61.6	1034.0	1560.0	626.1	6033.0	5744.2
P2	31.0	64.8	887.0	1337.1	590.1	6033.0	5541.0
P3	31.5	62.2	947.6	1505.1	626.6	6033.0	5692.5

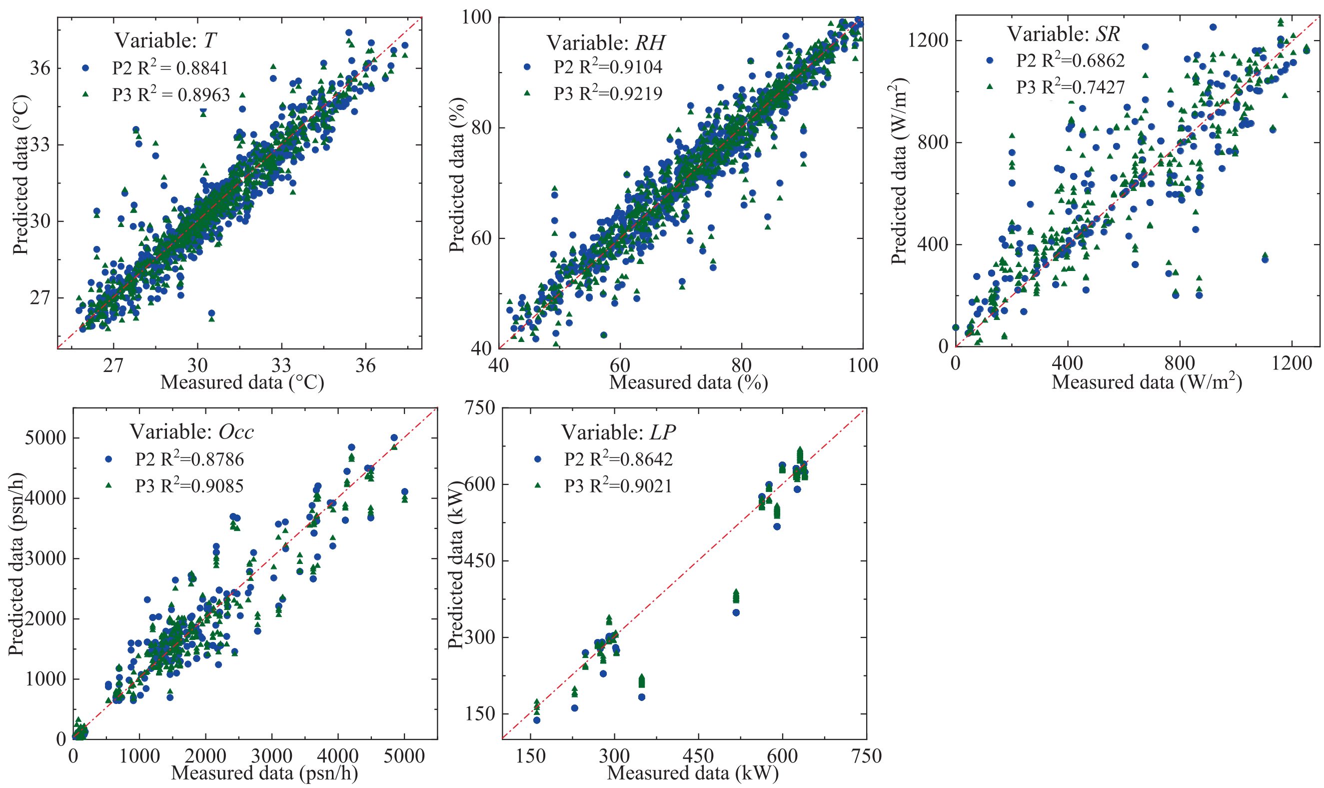

Fig. 11. Regression plot of P2 and P3 data types before and after calibrating.

Table 7
Evaluation of P2 and P3 data types.

Variables	Unit	Data type	RMSE	CV (%)	MAPE (%)
T	°C	P2	0.811	2.66	1.77
		P3	0.781	2.56	1.59
RH	%	P2	3.91	5.36	3.98
		P3	3.68	5.04	3.47
SR	W/m^2	P2	174.5	30.14	36.32
		P3	164.8	28.46	30.62
Occ	psn/h	P2	410.7	23	19.8
		P3	349.2	19.56	19.7
LP	kW	P2	60.12	11.20	9.57
		P3	48.63	9.06	7.4

CL_a(t) 5782 kW. The prediction relative error is −1.5% when using the calibrated input data P3, lower than relative error −4.2% when using the predicted input data P2. This provides that the uncertainty analysis of forecast input variables based on the MC method is beneficial to improve load prediction results reliability.

The database type P2 and P3 before and after calibrating are shown in Fig. 11. The evaluation performance of data types P2 and P3 compared to P1 is summarized in Table 7. Fig. 11 and Table 7 show that the uncertainty of all variables forecast data had been reduced with different degrees after calibration, especially *T* and *RH* were very close to the measured values. R^2 value of data type P3 was higher than P2. Three evaluation indexes of P3 was lower than P2 after calibrating. CV values of different variables were lower than 30% after calibrating. MAPE values of *T* and *RH* were 1.59% and 3.47% for P3, respectively, which were lower than 1.77% and 3.98% for P2. Results also indicate calibration with the MC method can reduce the uncertainties of input variables forecast data.

5.3. Load prediction results analysis

As uncertainty analysis and MC calibration method described in Sections 5.1 and 5.2, both prediction results using input variables P2 and 93 are compared with the on-site measured cooling load, as shown in Figs. 12 and 13. Although the prediction results of P2 and P3 were still different from the measured load at some points, the prediction results of P3 was closer to the measurements from the overall level than P2, e.g., R^2 value of P3 was 0.9627 higher than 0.9488 of P2. This demonstrates that uncertainty analysis of forecast input data for cooling load prediction can improve the accuracy of load prediction. Table 8 summarizes the evaluation indices of prediction performance using the predicted input data P2 and the calibrated data P3. Four evaluation indices of P3 is better than that of P2 and closer to P1. MAPE values of P2 and P3 are 16.6%, 14.3%, respectively. RMSE values of P2 and P3 are 472.6 kW and 381.1 kW, respectively, while RMSE value of P1 is 378.8 kW.

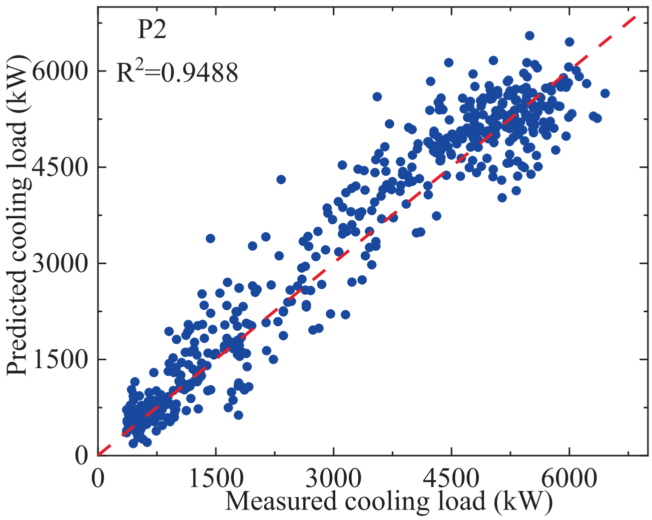

Fig. 12. The prediction results using the predicted input data P2.

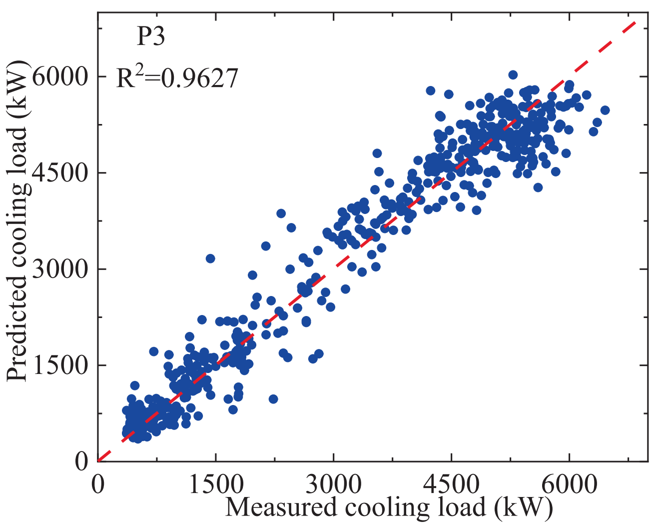

Fig. 13. The prediction results using the calibrated input data P3.

Besides, the CV value of P3 is less than 15%, which means that the model had good prediction results for engineering applications.

Fig. 14 shows a box-plot of different data types, which reflect the characteristics of the data distribution. It can be found that the prediction data distribution characteristics using P1 was closer to measured load data than using P2 and P3. The mean value of prediction results using P3 was closer to P2 after calibrating based on the MC method. The percentages of prediction error distribution within ±20% are 79.9%, 75.1%, and 81.1% for P1, P2, and P3, respectively. Comparing to P2, the prediction error of P3 can be reduced by 6%. This result indicates that the reliability of the predicted load was enhanced by calibrating the forecast input variables.

5.4. Discussion

This paper analyzes the effect of uncertainty of input variables on prediction reliability based on the MC method. Four evaluation indexes were used to compare the prediction results between different input data types, namely P1 (measured input data), P2 (forecast input data), and P3 (calibrated input data). It was found that the outdoor temperature *T* had the greatest sensitivity on the prediction result among the selected variables by sensitivity analysis.

Table 8
Evaluation indices of prediction performance of P2 and P3.

Input data type	R^2	RMSE (kW)	CV (%)	MAPE (%)
P2	0.9488	472.6	18.2	16.6
P3	0.9627	381.1	14.7	14.3

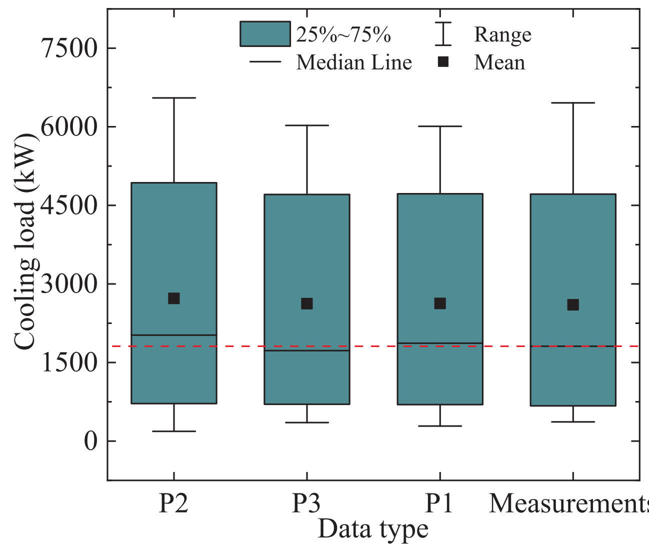

Fig. 14. The box-plot of the cooling load database.

The evaluation indices illustrated that the load prediction accuracy with P3 was higher than the load prediction accuracy using P2 directly, and closer to the load prediction accuracy using P1. Results also indicated that calibrating the input variable uncertainties was beneficial to reduce input uncertainties and can improve the reliability of load prediction.

The prediction results include two parts of uncertainties: one is the uncertainty from the forecast input variables, and the other is the uncertainty from the prediction model itself. The predicted input variables are random and uncertain, and it is difficult to completely guarantee accuracy. The method proposed in this paper can reduce the uncertainty of the predicted input variables, thereby reducing the prediction errors of the cooling load. However, this method also has its limitations. The error probability description cannot always adapt to the random conditions of outdoor weather, such as a sudden change in clouds leading to a sharp change in the solar radiation. Besides, the uncertainty from the prediction model itself cannot be eliminated even though SVM can effectively solve nonlinearities and has a strong generalization ability. The grid search algorithm in SVM takes more than 5-h to finish the entire prediction process. With the development of artificial intelligence algorithm technology, the optimal algorithm could reduce optimization time and uncertainty of the model itself as much as possible.

The current work is limited by the amount of historical indoor heat gains data and weather forecast data, so the highest accuracy of the error probability description cannot be completely achieved. To facilitate the application, it is necessary to improve the accuracy of the prior error probability description by collecting more measured data sources (i.e., weather data and historical cooling load). The larger the historical data collected, the more accurate the error probability distribution. The calibrated input data using this error probability will be closer to the measured data. Besides, the calibration process of input data based on MC method can not only be applied to the case building but also other types of buildings, e.g., office buildings and shopping markets. The calibrated cooling load can be used to guide the HVAC system optimal operation control. Examples include advanced air-conditioning control, chiller startup, and controlling the number of chilled water pumps. To integrate offline MC simulation and online load prediction, it still requires continuous efforts to develop a convenient module for practical application.

6. Conclusions

This paper mainly focuses on adopting the MC method to deal with the uncertainties of forecast input variables when conducting cooling load prediction. SVM model was used as the main model for cooling load prediction and various variables were considered. With sensitivity analysis, weather variables (i.e., outdoor dry-bulb temperature (*T*), relative humidity (*RH*), solar radiation (*SR*)), occupancy (*Occ*), and lighting power (*LP*) were selected as significant variables for cooling load prediction model. The prediction error was calculated using the difference between forecast input variables data and measured input data. The error characteristic was described with the specific probability distribution model. Sampling results of the specific probability were generated using a random sampling method. Then, these samplings were used to calibrate input variables forecast data.

A real library building located in Guangzhou was used to gather on-site measured load data for case studies. Case studies showed that the uncertainties of all input variables forecast data were reduced with different degrees after calibration, especially *T* and *RH* were very close to the measured values. Such as MAPE values of *T* and *RH* were 1.59% and 3.47% for P3, respectively, which were lower than 1.77% and 3.98% for P2. For all input variables, R^2 values of P3 were higher than P2. The uncertainty analysis results indicated the cooling load prediction accuracy of using the calibrated input data P3 (R^2 = 0.9627) was better than that of using the forecast input data P2 directly (R^2 = 0.9488), which was closer to load prediction results by using the real input data P1 (R^2 = 0.9632). The CV value using calibrated input data was less than 15%, which means that the model has good prediction results and is sufficiently suitable for engineering applications.

CRediT authorship contribution statement

Chengliang Fan: Writing - original draft, Methodology, Software, Validation. **Yundan Liao:** Writing - original draft, Methodology. **Guang Zhou:** Validation. **Xiaoqing Zhou:** Supervision, Project administration. **Yunfei Ding:** Resources.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was financially supported by the Guangzhou Housing and Urban-Rural Construction Committee (No. J-2016-11), the Key Project Fund for the Collaborative Innovation of the Universities from the Bureau of Education of Guangzhou City in China (No. 1201610004), the Key Projects of Basic Research and Applied Basic Research of Universities in Guangdong Province (No. 2018KZDXM050). The author would also like to acknowledge the support from the Guangzhou University Graduate Innovation Ability Development Funding Program (No. 2019GDJC-D16).

References

[1] B. Kim, Y. Yamaguchi, S. Kimura, Y. Ko, K. Ikeda, Y. Shimoda, Urban building energy modeling considering the heterogeneity of HVAC system stock: a case study on Japanese office building stock, Energy Build. 199 (2019) 547–561.
[2] R. Jing, M. Wang, R. Zhang, N. Li, Y. Zhao, A study on energy performance of 30 commercial office buildings in Hong Kong, Energy Build. 144 (2017) 117–128.
[3] C. Fan, Y. Ding, Cooling load prediction and optimal operation of HVAC systems using a multiple nonlinear regression model, Energy Build. 197 (2019) 7–17.
[4] Y. Sun, S. Wang, G. Huang, Model-based optimal start control strategy for multi-chiller plants in commercial buildings, Build. Serv. Eng. Res. Technol. 2 (2010) 113–129.
[5] X. Li, J. Wen, Review of building energy modeling for control and operation, Renew. Sustain. Energy Rev. 37 (2014) 517–537.
[6] D. Antonucci, U.F. Oberegger, W. Pasut, A. Gasparella, Building performance evaluation through a novel feature selection algorithm for automated arx model identification procedures, Energy Build. 150 (2017) 432–446.
[7] C. Fan, Y. Liao, Y. Ding, Development of a cooling load prediction model for air-conditioning system control of office buildings, Int. J. Low-Carbon Technol. 14 (1) (2019) 70–75.
[8] Y. Guo, E. Nazarian, J. Ko, K. Rajurkar, Hourly cooling load forecasting using time-indexed ARX models with two-stage weighted least squares regression, Energy Convers. Manage. 80 (2014) 46–53.
[9] R. Sarwar, H. Cho, S.J. Cox, P.J. Mago, R. Luck, Field validation study of a time and temperature indexed autoregressive with exogenous (ARX) model for building thermal load prediction, Energy 119 (2017) 483–496.
[10] G. Fu, Deep belief network based ensemble approach for cooling load forecasting of air-conditioning system, Energy 148 (2018) 269–282.
[11] S. Kumar, S.K. Pal, R.P. Singh, A novel method based on extreme learning machine to predict heating and cooling load through design and structural attributes, Energy Build. 176 (2018) 275–286.
[12] C. Xu, H. Chen, W. Xun, Z. Zhou, T. Liu, Y. Zeng, T. Ahmad, Modal decomposition based ensemble learning for ground source heat pump systems load forecasting, Energy Build. 194 (2019) 62–74.
[13] H. Zhong, J. Wang, H. Jia, Y. Mu, S. Lv, Vector field-based support vector regression for building energy consumption prediction, Appl. Energy 242 (2019) 403–414.
[14] C. Fan, Y. Ding, Y. Liao, Analysis of hourly cooling load prediction accuracy with data-mining approaches on different training time scales, Sustain. Cities Soc. 51 (2019).
[15] P. Eguía, E. Granada, J.M. Alonso, E. Arce, A. Saavedra, Weather datasets generated using kriging techniques to calibrate building thermal simulations with TRNSYS, J. Build. Eng. 7 (2016) 78–91.
[16] A.S. Anđelković, I. Mujan, S. Dakić, Experimental validation of a EnergyPlus model: application of a multi-storey naturally ventilated double skin façade, Energy Build. 118 (2016) 27–36.
[17] Y. Sun, S. Wang, F. Xiao, Development and validation of a simplified online cooling load prediction strategy for a super high-rise building in Hong Kong, Energy Convers. Manage. 68 (2013) 20–27.
[18] G. Qiang, T. Zhe, D. Yan, Z. Neng, An improved office building cooling load prediction model based on multivariable linear regression, Energy Build. 107 (2015) 445–455.
[19] Y. Guo, E. Nazarian, J. Ko, K. Rajurkar, Hourly cooling load forecasting using time-indexed ARX models with two-stage weighted least squares regression, Energy Convers. Manage. 80 (2014) 46–53.
[20] X. Li, J. Wen, Building energy consumption on-line forecasting using physics based system identification, Energy Build. 82 (2014) 1–2.
[21] J. Wang, G. Huang, Y. Sun, X. Liu, Event-driven optimization of complex HVAC systems, Energy Build. 133 (2016) 79–87.
[22] G. Huang, T.T. Chow, Uncertainty shift in robust predictive control design for application in CAV air-conditioning systems, Build. Serv. Eng. Res. Technol. 32 (4) (2011) 329–343.
[23] A.F. Janabi-Sharifi, Theory, and applications of HVAC control systems – a review of model predictive control (MPC), Build. Environ. 72 (2014) 343–355.
[24] Y. Liao, Uncertainty analysis for chiller sequencing control, Energy Build. 85 (2014) 187–198.
[25] Y. Liao, Y. Sun, G. Huang, Robustness analysis of chiller sequencing control, Energy Convers. Manage. 103 (2015) 180–190.
[26] W. Tian, Y. Heo, P. De Wilde, Z. Li, D. Yan, C.S. Park, X. Feng, G. Augenbroe, A review of uncertainty analysis in building energy assessment, Renew. Sustain. Energy Rev. 93 (2018) 285–301.
[27] M.M. Liu, Probabilistic prediction of green roof energy performance under parameter uncertainty, Energy 77 (2014) 667–674.
[28] A.S. Silva, E. Ghisi, Uncertainty analysis of the computer model in building performance simulation, Energy Build. 76 (2014) 258–269.
[29] A. Prada, F. Cappelletti, P. Baggio, A. Gasparella, On the effect of material uncertainties in envelope heat transfer simulations, Energy Build. 71 (2014) 53–60.
[30] Z. Li, G. Huang, Re-evaluation of building cooling load prediction models for use in humid subtropical area, Energy Build. 62 (2013) 442–449.
[31] P. Huang, G. Huang, G. Augenbroe, S. Li, Optimal configuration of multiple-chiller plants under cooling load uncertainty for different climate effects and building types, Energy Build. 158 (2018) 684–697.
[32] Y. Liao, G. Huang, Y. Ding, H. Wu, Z. Feng, Robustness enhancement for chiller sequencing control under uncertainty, Appl. Therm. Eng. 141 (2018) 811–818.

[33] Z. Liu, H. Tan, D. Luo, G. Yu, J. Li, Z. Li, Optimal chiller sequencing control in an office building considering the variation of chiller maximum cooling capacity, Energy Build. 140 (2017) 430–442.
[34] A. Parisio, M. Molinari, D. Varagnolo, K.H. Johansson, Energy management systems for intelligent buildings in smart grids, in: Intelligent Building Control Systems, Springer, Cham, 2018, pp. 253–291.
[35] M. Maasoumy, M. Razmara, M. Shahbakhti, A.S. Vincentelli, Handling model uncertainty in model predictive control for energy efficient buildings, Energy Build. 77 (2014) 377–392.
[36] P. Huang, Y. Wang, G. Huang, G. Augenbroe, Investigation of the ageing effect on chiller plant maximum cooling capacity using Bayesian Markov Chain Monte Carlo method, J. Build. Perform. Simul. 9 (5) (2016) 529–541.
[37] J. Zhao, Y. Duan, X. Liu, Uncertainty analysis of weather forecast data for cooling load forecasting based on the Monte Carlo method, Energies 11 (7) (2018) 1900.
[38] P. Huang, G. Huang, Y. Wang, HVAC system design under peak load prediction uncertainty using multiple-criterion decision making technique, Energy Build. 91 (2015) 26–36.
[39] B. Gunay, W. Shen, G. Newsham, Inverse blackbox modeling of the heating and cooling load in office buildings, Energy Build. 142 (2017) 200–210.
[40] Y. Ding, X. Liu, A comparative analysis of data-driven methods in building energy benchmarking, Energy Build. 209 (2020).
[41] F. Domínguez-Muñoz, J.M. Cejudo-López, A. Carrillo-Andrés, Uncertainty in peak cooling load calculations, Energy Build. 42 (7) (2010) 1010–1018.
[42] J. Zhao, X. Liu, A hybrid method of dynamic cooling and heating load forecasting for office buildings based on artificial intelligence and regression analysis, Energy Build. 174 (2018) 293–308.
[43] Q. Li, Q. Meng, J. Cai, H. Yoshino, A. Mochida, Applying support vector machine to predict hourly cooling load in the building, Appl. Energy 86 (10) (2009) 2249–2256.
[44] Y. Wei, X. Zhang, Y. Shi, L. Xia, S. Pan, J. Wu, M. Han, X. Zhao, A review of data-driven approaches for prediction and classification of building energy consumption, Renew. Sustain. Energy Rev. 82 (2018) 1027–1047.
[45] T.A. Reddy, I. Maor, C. Panjapornpon, Calibrating detailed building energy simulation programs with measured data-part II: application to three case study office buildings (RP-1051), HVAC&R Res. 13 (2) (2007) 243–265.